Noname manuscript No.
(will be inserted by the editor)

Misako Takayasu · Hayafumi Watanabe · Hideki Takayasu

Generalized central limit theorems for growth rate distribution of complex systems

Received: date / Accepted: date

Abstract We introduce a solvable model of randomly growing systems consisted of many independent subunits. Scaling relations and growth rate distributions in the limit of infinite subunits are analyzed theoretically. Various types of scaling properties and distributions reported for growth rates of complex systems in wide fields can be derived from this basic physical model. Statistical data of growth rates for about 1 million business firms are analyzed as an example of randomly growing systems in the real-world. Not only scaling relations are consistent with the theoretical solution, the whole functional form of the growth rate distribution is fitted with a theoretical distribution having a power law tail.

Keywords Central limit theorem · Growth rates · Stable distribution · Power laws · Firm statistics · Gibrat's laws

1 Introduction

Growth phenomena are generally highly irreversible dynamical processes far from thermal equilibrium [40]. From the viewpoint of statistical physics it is an important new topic that growth rates of complex systems often show non-trivial similar statistical behaviors across fields of sciences. Fat-tailed distributions of growth rates and a non-trivial shrink of variance as a function of the size are reported in various fields of sciences; business firms [1,6,8,32], sales of pharmaceutical products [10], circulation numbers of newspapers [23], population of migratory birds [15], animals' metabolic rate fluctuations [17], the amount of scientific funding [24], group size of religious activities [22], population size of cities [12], country's whole economic activity observed by GDP [18] and the amount of exports and governmental debts [25]. Probability densities of growth rates in most of these examples are typically approximated by double exponential (Laplace) distributions or by power law distributions, quite interestingly not by a Gaussian distribution.

Statistics on growth rates of business firms have a long history of study, and recently statistical physicists are involved in this topic. Gibrat postulated the "law of proportional effect" that the expected value of the growth rate of a business firm is proportional to the current size of the firm [14,25]. The original Gibrat's assumption states that the variance of growth rates is independent of the size, however,

Misako Takayasu · Hayafumi Watanabe*
Department of Computational Intelligence and Systems Science, Interdisciplinary Graduate School of Science and Engineering, Tokyo Institute of Technology, 4259-G3-52 Nagatsuta-cho, Midori-ku, Yokohama 226-8502, Japan

Hideki Takayasu
Sony Computer Science Laboratories Inc., 3-14-13 Higashigotanda, Shinagawa-ku, Tokyo 141-0022, Japan
Meiji Institute for Advanced Study of Mathematical Sciences, 1-1-1 Higashimita, Tama-ku, Kawasaki 214-8571, Japan
E-mail: h-watanabe@smp.dis.titech.ac.jp*

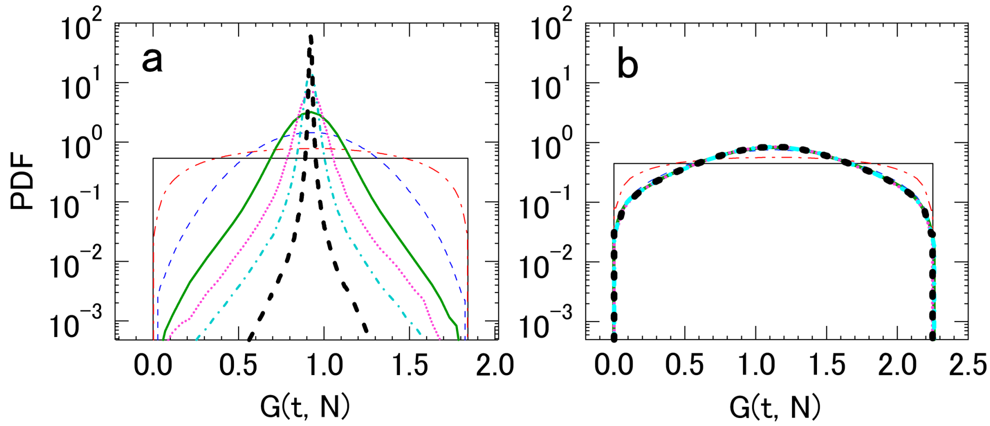

Fig. 1 Comparison of system size dependence of the probability density functions of growth rates for different values of α. a; The case of $\alpha = 1.5$ ($< g_j(t)^{1.5} >= 1$) and b; $\alpha = 0.5$ ($< g_j(t)^{0.5} >= 1$) in semi-log plots. In both cases the growth rate distribution of individual subunits follows a uniform distribution. $N = 1$ (black thin line), $N = 2$ (red dash-dotted line), $N = 10$ (blue broken line), $N = 10^2$ (bold green line), $N = 10^3$ (purple dotted line), $N = 10^4$ (light blue dash-dotted line) and $N = 10^6$ (black broken line).

data analyses of business firm activities show various types of variance-size relations. There are papers which support the original Gibrat's assumption [11,33], on the other hand non-trivial fractional power laws are reported not only for business firms but also in many other phenomena [1,6,8,10,12,15,17, 18,22,23,24,32,40]. Also, the country-dependence [21] and the transition from Gibrat's assumption to such power law decays are pointed out [37]. Various types of theoretical models of business firms have been introduced by physicists for better understanding of scaling properties of business firms from the standpoint of complex systems [2,3,4,7,20,26,27,31,39,41,42], however, there has been no unified theory which can explain all these basic properties simultaneously.

In the next section we introduce a new type of random growth model of a complex system consisted of many independent subunits, and in section 3 we theoretically solve non-trivial scaling relations between the variance of growth rates and the number of subunits by introducing a kind of renormalization. In the limit of infinite number of subunits the distribution of growth rates is shown to converge to a stable distribution with a power law tail. The stable distribution and the generalized central limit theorem were established in mathematics about 80 years ago [9,19], however, these concepts were applied mostly for theoretical models assuming scaling properties for phenomena such as turbulence [35, 36]. Validity of the theoretical results for growth rates is confirmed in section 4 by analyzing a huge database about business firms. This example may be the first real world application of an asymmetric stable distribution fitted for the whole scale range. The final section is devoted for discussion.

2 The model

We consider a system consisting of N subunits characterized by non-negative scalar quantities, $\{x_j(t)\}$. For each subunit we assume the following random multiplicative time evolution [16], which is known to be one of the basic process of producing power law fluctuations [29,38] . (See Appendix A for brief review of random multiplicative process).

$$x_j(t + \Delta t) = g_j(t)x_j(t) + f_j(t), \tag{1}$$

where $g_j(t)$ and $f_j(t)$ for $j = 1, 2, \cdots, N$, are growth rates and external forces, respectively, both are assumed to be independent identically distributed random variables taking only positive values. In the case that the probability of occurrence of $g_j(t) > 1$ is not zero and if $< \log(g_j(t)) >< 0$, where $< \cdots >$ denotes the average, it is known that there exists a statistically steady state in which the cumulative distribution follows a power law [16],

$$P(> x_j) \propto x_j^{-\alpha}, \tag{2}$$

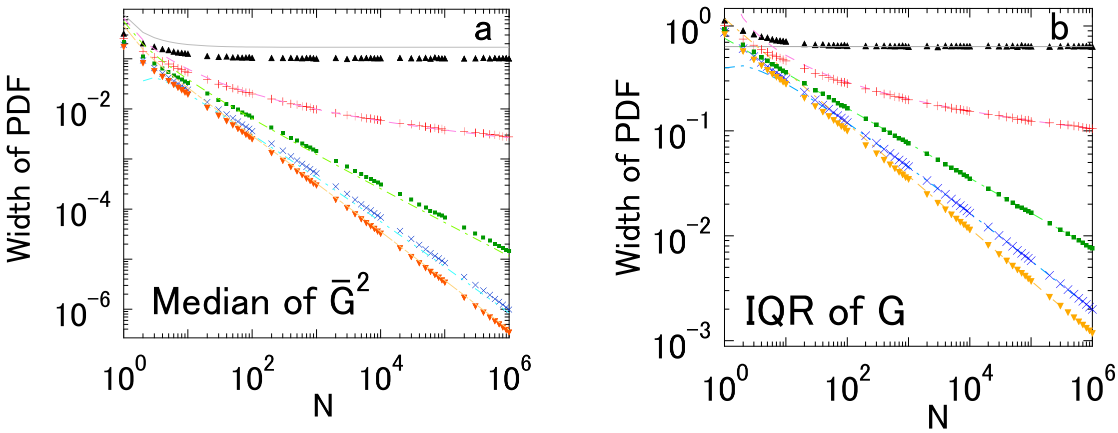

Fig. 2 Confirmation of theoretical estimates of size dependence of the width of the growth rate distribution. The width of growth rate distributions are measured by repeating numerical simulations for various values of N. The points are plotted in log-log scale for five cases: $\alpha = 0.5$, 1.0, 1.5, 2.0 and 2.5, (black, red, green, blue and orange), simulation results(points) and theory (lines). For numerical simulations, the growth rate for each subunit, $g_j(t)$, distributes uniformly in the range $(0, (\alpha+1)^{1/\alpha}]$ satisfying $< g_j(t)^\alpha >= 1$ ($\alpha = 0.5$, 1.0 ,1.5, 2.0 and 2.5). (a)Data shown are median of $\bar{G}^2$ for numerical simulations and Eq. 10 for theoretical estimations. (b)Data shown are the interquartile range (IQR) of G for numerical simulations and the square roots of the asymptotic functional forms of Eq. 10 given in Table 1 for theoretical estimations. The theoretical curves are shifted along the vertical axis so that they overlap with the numerical simulations at $N = 10^6$ in the panel (b). The IQR, which is the one of the most commonly-used robust measure of width of a PDF, is difference between the largest and smallest values in the middle 50% of a set of data. From these figures, we can confirm that the medians of $\bar{G}(t;N)^2$ almost correspond to the width given by Eq. 10, and IQRs are proportion to asymptotic behaviour of this equation given by table 1.

for large value of $x_j$with the positive exponent determined exactly only by the statistics of the growth rate by solving the following equation,

$$< g_j(t)^\alpha >= 1. \tag{3}$$

Based on a renormalization point of view we pay attention to the sum of all subunits, $X(t;N) \equiv \sum_{j=1}^{N} x_j(t)$, which follows the same type of time evolution as that of the subunits,

$$X(t + \Delta t; N) = G(t;N)X(t;N) + F(t;N), \tag{4}$$

where $\mathrm{F}(t;N) \equiv \sum_{j=1}^{N} f_j(t)$, and the growth rate of the whole system is defined as

$$G(t;N) \equiv \frac{\sum_{j=1}^{N} g_j(t)x_j(t)}{\sum_{j=1}^{N} x_j(t)}. \tag{5}$$

It is easy to show that the mean value of growth rates is invariant, namely, $< G(t;N) >=< g_j(t) >\equiv G$.

Properties of this model can be investigated by numerical simulation. Fig. 1a shows an example of deformation of growth rate distributions for various values of N in the case that Eq. (3) is fulfilled with $\alpha = 1.5$ observed in the statistically steady state realized for time steps larger than 10^6. Here, the distribution of growth rates of subunits is given by an independent uniform distribution as shown

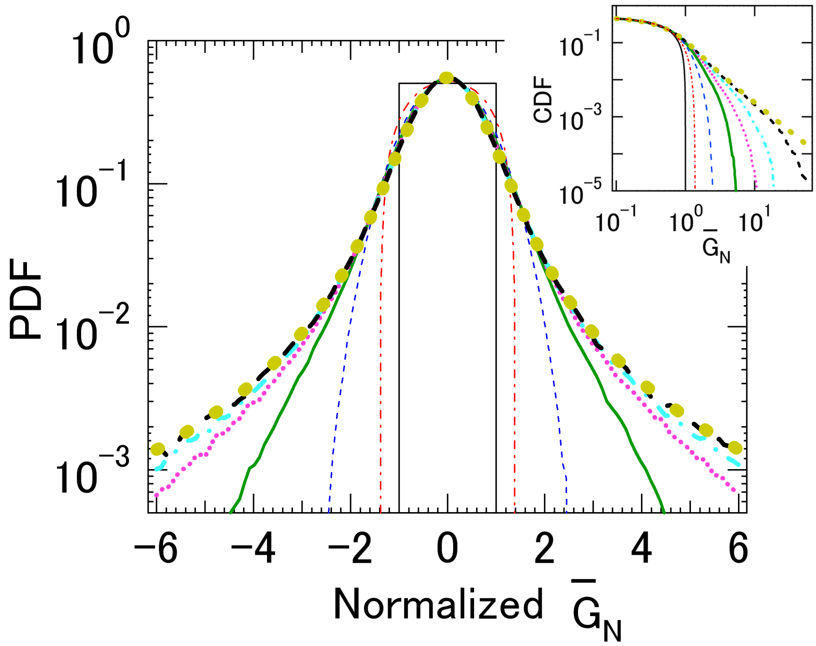

Fig. 3 Convergence of the normalized growth rate distributions in the case of $\alpha = 1.5$. The growth rate for each subunit, $g_j(t)$, distributes uniformly in the range $(0, (2.5)^{2/3}]$ satisfying $< g_j(t)^{1.5} >= 1$. The renormalized growth rate's probability density functions for systems with N subunits are plotted for several values of N. $N = 1$ (black thin line), $N = 2$ (red dash-dotted line), $N = 10$ (blue broken line), $N = 10^2$ (bold green line), $N = 10^3$ (purple dotted line), $N = 10^4$ (light blue dash-dotted line), $N = 10^6$ (black broken line), and the theoretical symmetric stable distribution, $p(\bar{G}; 1.5, 0)$ (yellow dotted line). The inserted figure shows the cumulative distribution function of the positive part in log-log scale confirming convergence to the power law.

by the case of $N = 1$. In this figure, the additive noise term, $f_j(t)$, is set to be a positive constant for simplicity as we confirmed that the main results do not depend on the values of $f_j(t)$ except the case that it is identically 0. As known from this figure the distribution of growth rates of the aggregated system tends to shrink slowly to a delta function as N goes to infinity. It is numerically confirmed that the same property of convergence to the delta-function holds for any distribution of growth rates of subunits if the growth rate distribution satisfies Eq. (3) with $\alpha \geq 1$.

Fig. 1b shows an example in the case of $\alpha = 0.5$. In this case the growth rate distribution stops shrinking for N larger than 10 and it converges to a non-trivial distribution in the limit of N goes to infinity. It is confirmed that this property is always observed if the value of α in Eq. (3) is between 0 and 1. The distribution of growth rate in the limit of $N = \infty$ depends on the functional form of the growth rate distribution for the subunits.

3 Theoretical analyses

We can theoretically evaluate the N-dependence of the width of the PDF of $G(t; n)$ by introducing an approximation of the random variables $\{x_j(t)\}$ which are known to follow a power law in the steady state. We introduce the following values as the measure of the width, $\bar{G}(t; N)^2$. Here, we define $\bar{G}(t; N)$ as:

$$\bar{G}(t; N) \equiv G(t; N) - G = \frac{\sum_{j=1}^{N} (g_j(t) - G) \cdot x_j(t)}{\sum_{j=1}^{N} x_j(t)}. \tag{6}$$

$\bar{G}(t; N)^2$ takes zero in the case that the PDF of $G(t; N)$ is a delta function. Let u be a random variable following a uniform distribution in the interval of $(0, 1]$. Then, the distribution of the new variable, $u^{-1/\alpha}$, follows a power law with exponent α. For uniform random variables, $\{u\}$, we can approximate

Value of α $< g_j(t)^\alpha >= 1$	Width-Size(N) relation (The limit of $N \to \infty$)	Limit distribution of growth rates
$2 < \alpha$ $< g_j(t)^2 >< 1$	$\frac{(\alpha-1)^2}{\alpha(\alpha-2)} N^{-1} \to 0$	Gaussian
$\alpha = 2$ $< g_j(t)^2 >= 1$	$\frac{\gamma+\log(N)}{4} N^{-1} \to 0$	
$1 < \alpha < 2$ $< g_j(t) >< 1$, $< g_j(t)^2 >> 1$	$\zeta(\frac{2}{\alpha}) \cdot (1-\frac{1}{\alpha})^2 \cdot N^{\frac{2}{\alpha}-2} \to 0$	Stable distribution with power law tails
$\alpha = 1$ $< g_j(t) >= 1$	$\frac{\zeta(2)}{(\gamma+\log(N))^2} \to 0$	$P(> \lvert\bar{G}\rvert \propto \lvert\bar{G}\rvert^{-\alpha})$
$0 < \alpha < 1$ $< g_j(t) >> 1$, $< \log(g(t)) >< 0$	$\frac{\zeta(2/\alpha)}{\zeta(1/\alpha)^2}$: constant (Gibrat's assumption holds)	Non-universal distribution Depending on the subunit's property

Table 1 Summary of generalized central limit theorems for growth rates. The value of α and the asymptotic functional form of Eq. 10 divided by σ^2, which corresponds to the width of the growth rate, for large system size, N, and the limit distributions. $\gamma = 0.577...$ is the Euler constant.

N samples by a set of deterministic values, $\{1/N, 2/N, \cdots, 1\}$, so that the set of power law variables $\{x_j(t)\}$, which follow Eq. (2), can be approximated by the deterministic set:
$\left\{N^{1/\alpha}, (N/2)^{1/\alpha}, (N/3)^{1/\alpha}, \cdots, (N/j)^{1/\alpha}, \cdots, 1\right\}$. Therefore, the typical sample of $\bar{G}^2(t;N)$ is obtained,

$$\bar{G_r}(t;N)^2 \equiv \frac{\{\sum_{j=1}^{N}(g_j(t)-G)\cdot(j/N)^{-1/\alpha}\}^2}{\left\{\sum_{j=1}^{N}(j/N)^{-1/\alpha}\right\}^2}. \tag{7}$$

Taking the average of $G_r(t;N)$ with respect to $g_j(t) (j = 1, 2, \cdots N)$ and applying the independence condition, $< (g_n(t) - G)(g_m(t) - G) >= \sigma^2\delta_{nm}$, we have

$$< \bar{G_r}(t;N)^2 >= \sigma^2 \cdot \frac{\sum_{j=1}^{N}(j/N)^{-2/\alpha}}{\left\{\sum_{j=1}^{N}(j/N)^{-1/\alpha}\right\}^2}, \tag{8}$$

where σ^2 is the variance of the growth rates for the subunits and δ_{nm} is Kroneker's delta. Then, we can calculate the summations in Eq. (8), $N^{1/\alpha}\sum_{j=1}^{N} j^{-1/\alpha}$ and $N^{2/\alpha}\sum_{j=1}^{N} j^{-2/\alpha}$, by applying an asymptotic expansion formula for the Riemann Zeta function,

$$\zeta(\lambda) \equiv \sum_{j=1}^{\infty}\frac{1}{j^\lambda} = \sum_{j=1}^{N}\frac{1}{j^\lambda} + \frac{1}{(\lambda-1)N^{\lambda-1}} - \frac{1}{2N^\lambda} + \cdots. \tag{9}$$

Neglecting the third term and higher order terms in the right hand side of Eq. (9), we have the following approximation of Eq. (8) for $0 < \alpha$:

$$\left\langle \bar{G_r}(t,N)^2 \right\rangle = \sigma^2 \cdot \frac{\zeta(\frac{2}{a}) - \frac{\alpha}{2-\alpha}N^{1-\frac{2}{\alpha}}}{\left\{\zeta(\frac{1}{\alpha}) - \frac{\alpha}{1-\alpha}N^{1-\frac{1}{\alpha}}\right\}^2}. \tag{10}$$

These theoretical evaluations are checked numerically in Fig. 2, in which the widths of growth rate distributions for the aggregated system are plotted as functions of the number of subunits, N. The theoretical asymptotic functional forms in Table 1 fit quite well asymptotically for all cases. It should be noted that the ordinary standard deviation or variance is not a good measure for the width of the distribution of growth rates as these quantities are affected easily by outliers. Instead of the standard deviation we introduce the median of $\bar{G}(t;N)^2$ and the interquartile range (IQR) for characterization as the measure of the width of distribution, in Figs. 2(a) and (b) respectively. The IQR, which is the one of the most commonly-used robust measure of width of a PDF, is difference between the largest and smallest values in the middle 50% of a set of data. From Fig. 2(a) we can confirm that the medians of $\bar{G}(t;N)^2$ almost correspond to the width given by Eq. 10, and from Fig. 2(b) IQRs are proportional to asymptotic behaviour of this equation given by table 1.

From Eq. (10) we find that the width of PDF of $G(t;N)$ converges to 0 following a power law of N in the case of $1 < \alpha$, while the width takes a finite value even in the limit of $N \to \infty$ in the case of $0 < \alpha < 1$. At the marginal case of $\alpha = 1$ the width decays to 0 very slowly for increasing N. For $2 < \alpha$ the width decays inversely proportional to N, that is, the typical N dependence in the case of ordinary central limit theorem. These functional forms of N dependence are summarized in Table 1 in the second column. As known from this result Gibrat's assumption of constant variance is fulfilled in the case of $0 < a < 1$, and the non-trivial power law decays of width-size relations observed in many complex systems are realized in the case of $1 < \alpha < 2$. As shown in the first column of this table the range of α is characterized by the form of equality or inequality for the moment function of the growth rates, $M(s) \equiv < g_j(t)^s >$. Derivation of these relations is summarized in Appendix B with the basic properties of the moment function. From this table we find that Gibrat's assumption holds for systems in which the mean growth rate is larger than 1, while the non-trivial power law decay of width-size relation for growth rate is expected in the situation that the mean growth rate of subunits is less than 1.

Next, we theoretically derive functional forms of the distribution of growth rates normalized by the width in the limit of $N \to \infty$. We consider the following 3 cases according to the value of α.

I: The case of $0 < \alpha < 1$*:* The width of growth rate does not shrink to 0 but it converges to a finite value even in the limit of $N \to \infty$ as known from Eq. (10). This reason can be understood by a general property of power law distribution with exponent less than 1. In such a case the mean value $< x_j(t) >$ diverges implying that some of samples in $\{x_j(t)\}$ take extraordinarily large values compared with others. So, both the denominator and numerator of Eq. (6) can be approximated by only finite numbers of extraordinarily large contributors, therefore, the value of Eq. (10) is finite even in the limit of large N. The distribution of growth rate shows the same property, namely, even in the limit of $N \to \infty$, the limit distribution is determined only by a small number of large contributors, therefore, we cannot expect a universal functional form in this case.

II. The case of $1 \le \alpha < 2$*:* As indicated in Fig.2 the width of distribution shrinks to 0 in the limit of $N \to \infty$ we can expect existence of a universal limit distribution independent of the initial condition. In this case the average, $< x_j(t) >$, takes a finite value in the steady state, the denominator in Eq. (6) can be roughly estimated as $N < x_j(t) >$ for very large values of N. On the other hand, the numerator in Eq. (6) is given by the sum of $(g_j(t) - G)x_j(t)$, in which $g_j(t) - G$ gives a coefficient taking either positive or negative sign randomly with respect to j, and $x_j(t)$ follows a power law with the exponent α. Namely, the numerator becomes a summation of N independent identically distributed random variables that have both positive and negative power law tails with the exponent α. As the generalized central limit theorem can be applied to such a sum of random variables, the limit distribution of growth rate $\tilde{G}$, which is normalized by the width of the distribution around the mean value, is expected to converge to a stable distribution, which has the form of an inverse Fourier transform [9], (see also Appendix C for a brief summary of the central limit theorem and the stable distributions),

$$p(\tilde{G};\alpha,\beta) \equiv \frac{1}{2\pi}\int_{-\infty}^{\infty} \exp\left\{-i\rho\tilde{G} - |\rho|^{\alpha}(1 - i\beta\psi(\rho,\alpha))\right\}d\rho, \tag{11}$$

where the asymmetry parameter β, which takes a value in the interval $[-1, 1]$, and the function $\psi(\rho, \alpha)$ are given as follows,

$$\beta \equiv \frac{< (g_j - G)|g_j - G|^{\alpha-1} >}{< |g_j - G|^{\alpha} >}, \quad \phi(\rho,\alpha) \equiv \frac{\rho}{|\rho|}\tan\frac{\pi\alpha}{2}. \tag{12}$$

It is well known that the limit probability density, Eq. (11), has a power law tail with the exponent α just like the distribution of $\{x_j(t)\}$.

In Fig. 3 we confirm the validity of this theoretical result by a numerical simulation for the case of $\alpha = 1.5$ and $\beta = 0$. The normalized growth rates for the system consisted of N-subunits, $\tilde{G}_N$, are calculated by subtracting the mean value and normalized by the width of the distribution. As known from this figure the distribution of normalized growth rates changes its functional form for different values of N. The probability density functions are converging clearly to the theoretical function, $p(\tilde{G}, 1.5, 0)$. It should be noted that this convergence is rather slow with respect to N and we can approximate

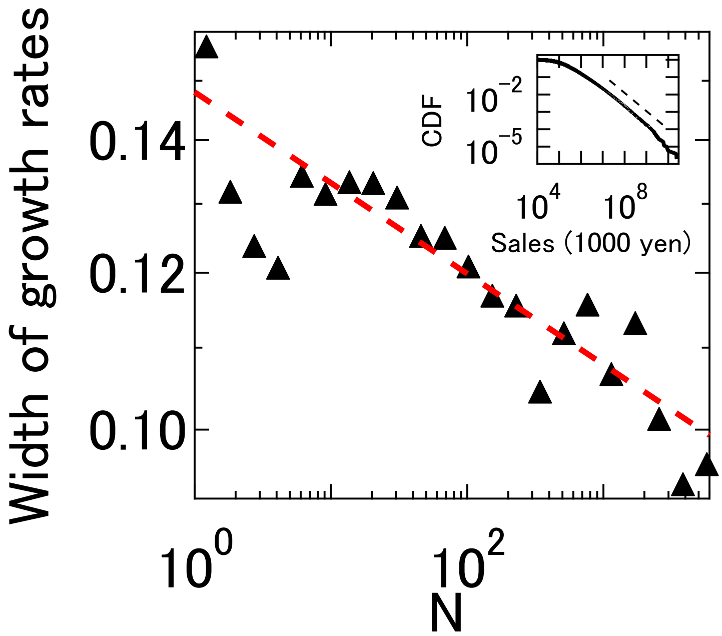

Fig. 4 Width of growth rate fluctuation of Japanese business firms as a function of number of workers, N, in log-log plot. The dashed line shows a theoretical curve given by the inverse power law, $N^{-0.046}$, which is derived from Eq. (10) with $\alpha = 1.05$. The inserted figure shows cumulative distribution function of annual sales of about 1 million Japanese firms in 2005 with the dotted line showing a power law distribution with the same exponent $\alpha = 1.05$.

the transient functional form typically by a double-exponential (Laplace) distribution as we can find for the case of $N = 10^2$. This slow convergence can be a reason for observation of double-exponential distributions in various observational reports [1,6,8,15,17,22,23,24,32].

III. The case of $2 \leq \alpha$*:* Similar estimation for denominator of Eq. (6) is valid and the ordinary central limit theorem can be applied to the numerator of Eq. (6) as the variances for $\{x_j(t)\}$ are finite. The expression of Eq. (11) is also valid in this case, however, the parameters are limited to $\alpha = 2$ and $\beta = 0$, namely, the limit distribution of normalized growth rate is always $p(\tilde{G}; 2, 0)$, which is the well-known Gaussian distribution with no long tail.

It is interesting to consider a special situation that the growth rates $\{g_j(t)\}$ are distributed symmetrically around 1. Then, we can derive $\alpha = 1$ from the basic relation $< g_j(t) >= 1$, and β in Eq. (12) is 0 by symmetry. In such a case we can expect that the limit distribution of normalized growth rate converges to $p(\tilde{G}; 1, 0) = 1/\left\{2\pi(1 + \tilde{G}^2)\right\}$ from Eq. (11).

Results for all these cases are summarized in the third column of Table 1. The limit distribution of growth rate is determined by the value of moment function for the growth rates of subunits. The important point is that the ordinary central limit theorem can be applied only in the limited cases of relatively small growth rates, the mean value of growth rate is less than 1 and the second moment of growth rate is less or equal to 1. For systems in the real world it is expected that the systems are nearly in the statistically steady state and the mean values of growth rates of subunits may take a value around 1. Then, as known from Table 1, the limit distributions of growth rates belong to either power laws or non-universal functional forms.

4 An application to business firm activities

Now, we apply these theoretical results for data analysis. Among various types of dynamical complex systems in the real world, business firms are attracting the attention of scientists because there are

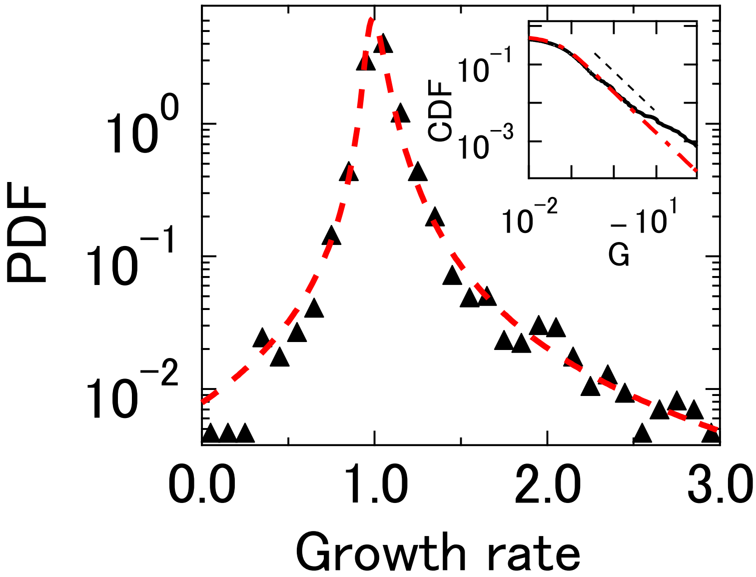

Fig. 5 Growth rate probability density function of large business firms with more than 300 workers (black triangles) in semi-log plot. The red dotted line shows the theoretical function of an asymmetric stable distribution given by $p(\tilde{G}; 1.05, 0.45)$ in Eq. (11) which is translated and re-scaled to fit the peak value and the width. The inserted figure shows the cumulative distribution function of the normalized positive growth rates (black line) in log-log plot to confirm the fitness with a theoretical stable distribution (red dashed line) and with a power law with tail exponent $\alpha = 1.05$ (black dotted line).

precise observation data in the form of financial reports [1,6,8,11,12,17,21,33]. In order to check the validity of our theory, we analyze comprehensive business firm data of Japan provided by the governmental research institute, RIETI (Research Institute for Economy, Trade and Industry). The data contains financial reports of 961,318 business firms, which practically covers all active firms in Japan in 2004 and 2005. It is already confirmed that basic quantities of these business firms, such as annual sales, incomes and number of employees, follow power law distributions [13].

There are several quantities that characterize the size of a business firm, such as assets, number of employees, sales, and incomes. Among these quantities, we focus on sales because this quantity reflects the present scale of activity of a firm most directly. Also, we simply assume that the whole activity of a business firm is given by the sum of the activities of individual employees; namely, we regard $X(t; N)$ in Eq. (4) as the annual sales of a business firm with N employees in the t-th year.

Neglecting both the additive term in Eq. (4) and the change of number of employees in a year, we estimate the growth rate $G(t; N)$ by the ratio of $t + 1$-th year's annual sale over the t-th year's sale of a business firm with N employees. It is already confirmed from the data that the autocorrelation of growth rate averaged over all business firms is very close to 0 implying that the growth of a business firm can be roughly viewed as an independent random growth process approximated by Eq. (4) with negligibly small external force term.

Categorizing the business firms by the number of employees, N, we can measure the width of growth rate distribution for each category. Fig. 4 shows N-dependence of the width of growth rate distribution in log-log scale. It is confirmed that Gibrat's assumption of size-independence does not hold in this case, and the width of growth rate decays clearly for large N. Here, the theoretical line is given by a power law, $N^{-0.046}$, which is derived from Eq. (10) in the case of $\alpha = 1.05$. In the inserted figure the cumulative distribution of annual sales of all firms, corresponding to a superposition of distribution of $X(t; N)$ for all N, is plotted in log-log scale. We can confirm that the tail of the distribution is approximated by a power law with the exponent $\alpha = 1.05$ as expected.

A theoretical limit distribution of growth rate, $p(\tilde{G}; 1.05, 0.45)$ in Eq. (11), is plotted together with that for real data estimated for N larger than 300 in Fig. 5. The inserted figure shows the cumulative

distribution of the normalized growth rate for positive $\tilde{G}$ in log-log scale to confirm the functional form of the fat-tail. The growth rate distribution is asymmetric in this case and the whole functional form is well approximated by the theoretical curve of the stable distribution. This may be the first real-world example of application of an asymmetric stable distribution with a fractional value of characteristic exponent fitted in the whole range since the birth of mathematical theory in the 1930's.

5 Discussions

In this paper we introduced a theoretically solvable model of sum of randomly growing independent subunits. As summarized in Table 1 we found generalized central limit theorems applicable for a composite of randomly growing subunits, in which we can find various types in both width-size relations and the limit distributions of growth rates. As an example of a real world application, we analyzed a huge database of business firm growth rates of Japan and confirmed consistency with the theory.

Application study of this theory for growth rate distribution of complex systems is highly encouraged. It is expected that the shrinking of growth rate width and the functional form of limit distributions can be directly compared with real data of various phenomena to check the universality of this aggregated system of randomly growing subunits. As already commented for Fig. 1, the normalized growth rate distribution looks quite similar to a double-exponential (Laplace) distribution in the case of intermediate numbers of subunits. This type of finite size effect should be treated carefully in real-world data analysis. There is the possibility that the varieties of empirically known properties of the growth rate of complex systems introduced in the beginning of this paper can be understood using our approach as a frame work.

Real-world systems may not be in a steady-state, so it is important for future work to investigate transient behaviors of this independent subunit system before it reaches the statistically steady state. Extension of this novel renormalization view of growth rates to cases of interacting subunits may also be an attractive new research topic. It is expected that a variation of generalized central limit theorem for the growth rates might be found for the wider category of growing complex systems like the case of non-standard statistical physics for long-range interaction systems [28].

Acknowledgements The authors appreciate RIETI for providing the business firm data. This work was partly supported by Research Foundations of the Japan Society for the Promotion of Science, Project Number 22656025 (M. T.) and Grant-in-Aid for JSPS Fellows no. 219685 (H. W.).

Appendix A: An introduction to random multiplicative process

As random multiplicative process is not widely known, here we introduce a simple exactly solvable case of random multiplicative process and show intuitively how the process realizes a power law distribution in the statistically steady state. Also, a continuum limit version of this multiplicative process is discussed.

We consider a positive random variable, $x(t)$, which follows the following stochastic equation,

$$x(t+\Delta t) = g(t)x(t) + 1, \tag{A.5}$$

where $g(t)$ is a stochastic noise term which takes either a positive constant, g, or 0 with probability 1/2, respectively. Starting with the initial condition, $x(0) = 1$, these time evolutions are given as

$$x(\Delta t) = \begin{cases} g+1 & (prob. \quad 1/2) \\ 1 & (prob. \quad 1/2) \end{cases}, x(2\Delta t) = \begin{cases} g^2+g+1 & (prob. \quad 1/4) \\ g+1 & (prob. \quad 1/4) \\ 1 & (prob. \quad 1/2) \end{cases}, \cdots \quad . \tag{A.5}$$

The general solution at time step τ is obtained as

$$x(\tau\Delta t) = \begin{cases} (g^\tau-1)/(g-1) & (prob. \quad 1/2^\tau) \\ (g^{\tau-1}-1)/(g-1) & (prob. \quad 1/2^\tau) \\ \vdots & \\ (g^k-1)/(g-1) & (prob. \quad 1/2^k) \\ \vdots & \\ 1 & (prob. \quad 1/2) \end{cases}, \tag{A.5}$$

where k is an integer from 1 to τ. From this solution we can calculate the cumulative distribution of $x(t)$ in the limit of $t \to \infty$ as

$$P(\geq x) = 2\{1 + (g-1)x\}^{-\frac{\log(2)}{\log(g)}}, \tag{A.5}$$

where $P(\geq x)$ denotes the probability that $x(\infty)$ takes a value larger than or equal to x. In the case that $g > 1$ we have the asymptotic power law distribution for very large value of x,

$$P(> x) \propto x^{-\alpha}, \tag{A.5}$$

where the exponent, $\alpha = \log(2)/\log(g)$, fulfills Eq. (3) in the range, $0 < \alpha < \infty$, namely,

$$< g(t)^{\alpha} > = \frac{1}{2} \cdot g^{\alpha} + \frac{1}{2} \cdot 0 = 1 \tag{A.5}$$

With this special example we confirm the validity of Eq. (1) to Eq. (3). Note that in this example, the stationary condition, $< \log(g(t)) > < 0$, is automatically satisfied because the condition that $g(t) = 0$ with probability 1/2 gives the value, $< \log(g(t)) > = -\infty$.

The key point of realizing the power law in this multiplicative random process is understood intuitively by neglecting the additive term. The probability of repeating $g(t) = g$ for k time steps is given as $p(k) \equiv 1/2^k = e^{-k\log(2)}$ and the corresponding value of $x(t)$ is approximated as $x(t) \approx g^k = e^{k\log(g)}$, then by deleting k from these relations we have Eq. (A.5). Namely, successive exponential growth with an exponential distribution of duration time gives the power law distribution.

This type of derivation of power law can be generalized in the following way. Let us consider the following general form of random multiplicative process,

$$x(t + \Delta t) = g(t)x(t) + f(t) \tag{A.5}$$

where $g(t)$ and $f(t)$ are independent random variables taking positive values, and we assume the situation that $\log g(t)$ fluctuates around 0 [30]. Taking logarithm for both sides and introducing variables $y(t) \equiv \log(x(t))$ and $r(t) = \log g(t)$, Eq. (A.5) can be transformed as

$$y(t + \Delta t) = \log\{g(t)x(t) + f(t)\} = y(t) + r(t) + \frac{f(t)}{g(t)x(t)} + \cdots \tag{A.5}$$

Neglecting the terms including $f(t)$ as higher order terms, time evolution of the probability density of $y(t)$, $p(y,t)$, is approximated by a Fokker-Plank equation.

$$p(y, t + \Delta t) \approx \int_{-\infty}^{\infty} \omega(r)p(y - r, t)dr = p(y,t) - < r > \frac{\partial p(r,t)}{\partial y} + \frac{< r^2 >}{2}\frac{\partial^2 p(r,t)}{\partial^2 y} + \cdots, \tag{A.5}$$

where $\omega(r)$ denotes the probability density of r. Assuming existence of a statistically steady state we have the following exponential distribution.

$$p(y) \propto \exp^{\frac{2<r>}{<r^2>}y} \tag{A.5}$$

In the situation, $< r > = < \log(g) > < 0$, which is equivalent to the condition of existence of steady state for random multiplicative process [16], Eq. (A.5) is shown to be equivalent to the power law, Eq. (A.5), with its exponent given as

$$\alpha \approx -\frac{2 < r >}{< r^2 >} \tag{A.5}$$

This value is derived from the exact relation, Eq. (3). By expanding the left hand side of the equation, $< g(t)^{\alpha} > = 1$, in terms α as follows, and by equating the second and third terms in right hand side as an approximation.

$$\begin{aligned} < g(t)^{\alpha} > &= < e^{\alpha \log(g(t))} > \\ &= 1 + \alpha < \log(g(t)) > + \alpha^2 \frac{< \{\log(g(t))^2\} >}{2} + \cdots. \end{aligned} \tag{A.5}$$

The key equation for determining the value of exponent, Eq. (3), can be derived roughly by the following way. Neglecting the additive term in the right hand side of Eq. (A.5) and taking an average over realizations after taking the s-th power of both sides, we have the following relation.

$$< x(t + \Delta t)^s > \approx < g(t)^s > < x(t)^s >, \tag{A.5}$$

In the case that $< g(t)^s >> 1$ it is clear that $< x(t)^s >$ diverges in the limit of $t \to \infty$. On the other hand in the case that $< g(t)^s > < 1$ the value of $< x(t)^s >$ is always finite. Therefore, we have the following relations,

$$< x(\infty)^s > = \begin{cases} \infty & (s > \alpha) \\ finite & (s < \alpha) \end{cases}, \tag{A.5}$$

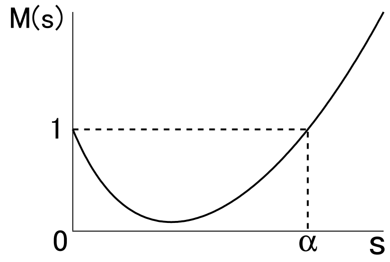

Fig. B-1 A schematic graph of the moment function.

where α satisfies Eq. (3). The property of Eq. (A.5) is one of typical characteristics of the power law distribution, Eq. (A.5). So, we can find that the power law exponent Eq. (A.5) is consistent with Eq. (3).

A rigorous mathematical derivation of this relation was done by Kesten in 1973 in more general form of real-valued matrix considering also the case that the distribution of the additive noise follows a power law [16]. In his proof the value of α is limited in the range of $0 < \alpha \leq 2$ as he applies the theory of stable distribution, however, our numerical analysis and the above intuitive theoretical analysis suggests that the value of α can be extended to the whole range $0 < \alpha$.

It should be noted that the existence of the additive term in Eq. (1) or Eq. (A.5) is essential to realize the statistically steady state. As known from Eq. (A.5) the stochastic process is a random walk with a negative trend in view of $\log(x(t))$, therefore, without any additive noise term the random walker tends to shrink to $x(t) = 0$, which is not the power law steady state. Even without the additive term ($f(t) \equiv 0$) the same power law steady state can also be realized by introducing a repulsive boundary condition such as requiring $x(t) \geq 1$ for Eq. (A.5) by adding a rule that $x(t + \Delta t) = 1$ when $x(t) < 1$.

Dividing both sides of Eq. (A.5) by Δt and considering the continuum limit of $\Delta t \to 0$, we have the following Langevin equation with a time-dependent viscosity.

$$\frac{d}{dt}X(t) = -\mu(t)X(t) + F(t), \tag{A.5}$$

where

$$X(t) \equiv x(t), \quad \mu(t) \equiv \lim_{\Delta t \to \infty} \frac{1 - g(t)}{\Delta t}, \quad F(t) \equiv \lim_{\Delta t \to \infty} \frac{f(t)}{\Delta t}. \tag{A.5}$$

As known from this equation the case of $g(t) > 1$ corresponds to a negative value of viscosity, $\mu(t) < 0$. In the case of a colloidal particle's diffusion in water such a negative viscosity cannot be realized, however, in the case of voltage fluctuation in an electric circuit, which is approximated also by a Langevin equation, we can consider a negative viscosity state by introducing an amplifier in the circuit. Namely, the value of $\mu(t)$ corresponds to resistivity in the electric circuit and in the situation that fluctuation of voltage is amplified as a whole, and the effective resistivity takes a negative value. By introducing an electric circuit in which an amplifier works at random timing we have a physical situation which is described by Eq. (A.5) and power law distributions of voltage fluctuation are confirmed experimentally [28].

Appendix B: Basic properties of the moment function

As Eq. (3) is the key of determining the exponent of the power law of Eq. (2), α, here, we summarize the basic properties of the moment function for growth rate of a subunit, $M(s) \equiv < g_j(t)^s >$. This is a continuous function and it is concave with respect to s for any distribution of $g_j(t)$ because the second derivative of this function is always positive, $M(s)'' \equiv < (\log(g_j(t))^2 g_j(t)^s > \quad > 0$. As $M(0) = 1$ is an identity, if $M(\alpha) = 1$ holds for a positive value of α, then we know that $M(s) < 1$ for $0 < s < \alpha$ and $M(s) > 1$ for $\alpha < s$ as schematically shown in Fig. B-1. So, $M(2) < 1$ corresponds to $2 < \alpha$ as shown in the first column of Table 1. In the situation that Eq. (3) holds with $0 < \alpha < 1$, then we have $M(1) =< g_j(t) >> 1$, while in the situation, $1 < \alpha$, we have $M(1) =< g_j(t) >< 1$.

The stationary condition, $< \log(g(t)) >< 0$, means that the slope of at the origin, $M(0)'$, is negative, so if this condition is not fulfilled $M(s) > 1$ for any positive s, implying that the stochastic process of Eq. (1) is not stationary. On the other hand in the case that the probability of occurrence of $g_j(t) > 1$ is 0, it is trivial that $M(s) < 1$ for any positive s.

Appendix C: A brief review of the generalized central limit theorem

The central limit theorem is one of the most powerful mathematical tools; however, it is too often used to approximate the sum of random variables, of the form, $Y(N) \equiv y_1 + y_2 + \cdots + y_N$, by normal distributions. In

fact there are three required conditions on random variables for their sums to obey the central limit theorem [9]:

1. All random variables follow the identical distribution.
2. The variables are independent.
3. The variance of the variables is finite.

If one of these conditions is violated, then the central limit theorem does not apply.

Violation of the first condition has recently been attracting attention as super-statistics, that is, superposition of stochastic variables having different statistics [5]. As an old example, a fat-tailed velocity distribution observed in randomly stirred granular particles can be explained by superposition of normal distributions having different variances due to clustering caused by inelastic collisions [34].

Giving a general discussion of correlated variables violating the second condition is rather difficult, as the details depend on the details of correlation. It is known that universal properties independent of the details of the system can be expected at the critical point of phase transition at which power law distributions and power law scaling relations play important roles [31]. Also, theoretical approaches based on the concept of "non-extensive entropy" can provide general solutions for strongly correlated systems having scale-free interactions such as charged particles [13]. However, considering activity of a business firm, for example, it seems very difficult to describe both internal and external interactions by a general mathematical formulation.

Violation of the third condition, infinite variance, was intensively studied in the 1930's by the pioneering mathematician P. Levy as "stable distributions" [19]. Assuming that $\{y_1, y_2, \cdots, y_N\}$ are independent identically distributed random variables, he showed that the fluctuation width of the sum $Y(N)$ increases proportional to $N^{1/\alpha}$ in general where α is called the characteristic exponent which lies in the range, $0 < \alpha \leq 2$. The limit distribution is defined for the normalized variable, $z \equiv \{Y(N) - c_N\}N^{-1/\alpha}$, where c_N is a term corresponding to the mean value. In the limit of $N \to \infty$, the distribution of z becomes a stable distribution that has a power law tail $p(z) \propto z^{-\alpha-1}$ for $0 < \alpha < 2$. The limit distribution converges to the normal distribution when $\alpha = 2$, according to the ordinary central limit theorem. This general result is called the generalized central limit theorem. The general functional form of the stable distribution is given by Eqs. (11) and (12) [9].

References

1. Amaral, L., Buldyrev, S., Havlin, S., Leschhorn, H., Maass, P., Salinger, M., Eugene Stanley, H., Stanley, M.: Scaling behavior in economics: I. empirical results for company growth. Journal de Physique I **7**, 621–633 (1997)
2. Amaral, L., Buldyrev, S., Havlin, S., Salinger, M., Stanley, H.: Power law scaling for a system of interacting units with complex internal structure. Phys. Rev.Lett. **80**, 1385–1388 (1998)
3. Aoki, M., Yoshikawa, H.: Reconstructing Macroeconomics: A Perspective from Statistical Physics and Combinatorial Stochastic Processes (Japan-US Center UFJ Bank Monographs on International Financial Markets). Cambridge University Press, Cambridge, Cambridge (2011)
4. Aoyama, H., Fujiwara, Y., Ikeda, Y., Iyetomi, H., Souma, W.: Econophysics and Companies: Statistical Life and Death in Complex Business Networks. Cambridge University Press, Cambridge (2011)
5. Beck, C., Cohen, E.: Superstatistics. Physica A: Statistical Mechanics and its Applications **322**, 267–275 (2003)
6. Bottazzi, G., Dosi, G., Lippi, M., Pammolli, F., Riccaboni, M.: Innovation and corporate growth in the evolution of the drug industry. International Journal of Industrial Organization **19**, 1161–1187 (2001)
7. Buldyrev, S., Growiec, J., Pammolli, F., Riccaboni, M., Stanley, H.: The growth of business firms: Facts and theory. J. Eur. Econ.Assoc. **5**, 574–584 (2007)
8. De Fabritiis, G., Pammolli, F., Riccaboni, M.: On size and growth of business firms. Physica A **324**, 38–44 (2003)
9. Feller, W.: An Introduction to Probability Theory and Its Applications Vol. 1, Edition 3, volume 1 edn. Wiley, New York, New York (1968)
10. Fu, D., Pammolli, F., Buldyrev, S., Riccaboni, M., Matia, K., Yamasaki, K., Stanley, H.: The growth of business firms: Theoretical framework and empirical evidence. Proc. Nat. Acad. Sci. USA **102**, 18,801–18,806 (2005)
11. Fujiwara, Y., Aoyama, H., Di Guilmi, C., Souma, W., Gallegati, M.: Gibrat and Pareto–Zipf revisited with european firms. Physica A **344**, 112–116 (2004)
12. Gabaix, X.: Zipf's law for cities: an explanation. Q. J. Econ. **114**, 739–767 (1999)
13. Gell-Mann, M., Tsallis, C.: Nonextensive entropy: interdisciplinary applications. Oxford University Press, USA, USA (2004)
14. Kalecki, M.: On the Gibrat Distribution. Econometrica **13**, 161–170 (1945)
15. Keitt, T., Stanley, H.: Dynamics of north american breeding bird populations. Nature **393**, 257–260 (1998)
16. Kesten, H.: Random difference equations and renewal theory for products of random matrices. Acta Mathematica **131**, 207–248 (1973)
17. Labra, F., Marquet, P., Bozinovic, F.: Scaling metabolic rate fluctuations. Proc. Nat. Acad. Sci. USA **104**, 10,900–10,903 (2007)
18. Lee, Y., Nunes Amaral, L., Canning, D., Meyer, M., Stanley, H.: Universal features in the growth dynamics of complex organizations. Phys. Rev. Lett. **81**, 3275–3278 (1998)
19. Lévy, P., Borel, M.: Théorie de l'addition des variables aléatoires, vol. 1. Gauthier-Villars, Paris, Paris (1954)

20. Miura, W., Takayasu, H., Takayasu, M.: Effect of coagulation of nodes in an evolving complex network. Phys. Rev. Lett. **108**, 168,701 (2012)
21. Okuyama, K., Takayasu, M., Takayasu, H.: Zipf's law in income distribution of companies. Physica A **269**, 125–131 (1999)
22. Picoli Jr, S., Mendes, R.: Universal features in the growth dynamics of religious activities. Phys. Rev. E **77**, 036,105 (2008)
23. Picoli Jr, S., Mendes, R., Malacarne, L., Lenzi, E.: Scaling behavior in the dynamics of citations to scientific journals. Europhys. Lett. **75**, 673 (2007)
24. Plerou, V., Amaral, L., Gopikrishnan, P., Meyer, M., Stanley, H.: Similarities between the growth dynamics of university research and of competitive economic activities. Nature **400**, 433–437 (1999)
25. Podobnik, B., Horvatic, D., Pammolli, F., Wang, F., Stanley, H., Grosse, I.: Size-dependent standard deviation for growth rates: Empirical results and theoretical modeling. Phys. Rev. E **77**, 056,102 (2008)
26. Riccaboni, M., Pammolli, F., Buldyrev, S., Ponta, L., Stanley, H.: The size variance relationship of business firm growth rates. Proc. Nat. Acad. Sci. USA **105**(50), 19,595–19,600 (2008)
27. Saichev, A., Malevergne, Y., Sornette, D.: Theory of Zipf's Law and Beyond (Lecture Notes in Economics and Mathematical Systems). Springer, Heidelberg, Heidelberg (2009)
28. Sato, A., Takayasu, H., Sawada, Y.: Power law fluctuation generator based on analog electrical circuit. Fractals **8**, 219–225 (2000)
29. Sornette, D.: Critical Phenomena in Natural Sciences: Chaos, Fractals, Selforganization And Disorder : Concepts And Tools (Springer Series in Synergetics), 2 edn. Springer, Berlin, Berlin (2006)
30. Sornette, D., Cont, R.: Convergent multiplicative processes repelled from zero: power laws and truncated power laws. Journal de Physique I **7**(3), 431–444 (1997)
31. Stanley, H.E.: Introduction to Phase Transitions and Critical Phenomena (International Series of Monographs on Physics), reprint edn. Oxford Univ Pr on Demand, Oxford (1987)
32. Stanley, M., Amaral, L., Buldyrev, S., Havlin, S., Leschhorn, H., Maass, P., Salinger, M., Stanley, H.: Scaling behaviour in the growth of companies. Nature **379**, 804–806 (1996)
33. Sutton, J.: Gibrat's legacy. Journal of economic Literature **35**, 40–59 (1997)
34. Taguchi, Y., Takayasu, H.: Power law velocity fluctuations due to inelastic collisions in numerically simulated vibrated bed of powder. Europhys. Lett. **30**, 499 (2007)
35. Takayasu, H.: Stable distribution and levy process in fractal turbulence. Progr. Theoret. Phys **72**, 471–479 (1984)
36. Takayasu, H.: $f^{-\beta}$Power Spectrum and Stable Distribution. J. Phys. Soc. Japan **56**, 1257–1260 (1987)
37. Takayasu, H., Okuyama, K.: Country dependence on company size distributions and a numerical model based on competition and cooperation. Fractals **6**, 67–79 (1998)
38. Takayasu, H., Sato, A., Takayasu, M.: Stable infinite variance fluctuations in randomly amplified langevin systems. Phys. Rev. Lett. **79**, 966–969 (1997)
39. Tamura, K., Miura, W., Takayasu, M., Takayasu, H., Kitajima, S., Goto, H.: Estimation of flux between interacting nodes on huge inter-firm networks. In: International Journal of Modern Physics: Conference Series, vol. 16, pp. 93–104. World Scientific (2012)
40. Vicsek, T.: Fractal Growth Phenomena: 1st Edition. World Scientific Publishing Company, Singapore, Singapore (1989)
41. Watanabe, H., Takayasu, H., Takayasu, M.: Biased diffusion on the japanese inter-firm trading network: estimation of sales from the network structure. New J. Phys. **14**(4), 043,034 (2012)
42. Yamasaki, K., Matia, K., Buldyrev, S., Fu, D., Pammolli, F., Riccaboni, M., Stanley, H.: Preferential attachment and growth dynamics in complex systems. Phys. Rev. E **74**, 035,103 (2006)